STIC Scientific and Technical Information Center	Dialog®

Core FT1:

Business & Industry , File 9 (1994 - present)

ABI/INFORM®, File 15 (1971 - present)

Gale Group PROMT®, File 16 (1990 - present)

Gale Group PROMT®, File 160 (1972-1989)

Gale Group Trade & Industry Database , File 148 (1976 - present)

Gale Group Computer Database , File 275 (full-text 1/1988 - present)

Business Wire, File 610 (Mar 1999 - present)

Business Wire, File 810 (1986 - February 1999)

Core FT2:

Dialog Global Reporter, File 20 (May 1997 - present)

The McGraw-Hill Companies Publications Online, File 624 (1985 - present)

Gale Group New Product Announcements/Plus® (NPA/Plus, File 621 (1985 - present)

Gale Group Newsletter Database , File 636 (1988 - present)

PR Newswire, File 613 (May 1999 - present)

San Jose Mercury News, File 634 (Jun 1985 - present)

PR Newswire, File 813 (May 1987 - May 1999)

Sub35FT:

McClatchy-Tribune Information Service, File 608 (Jan 1989 - present)

American Banker Financial Publications, File 625 (1981 - June 2008)

Banking Information Source, File 268 full-text (1994 - present)

Bond Buyer Full Text, File 626 (November 1981 - April 2008)

DIALOG Finance & Banking Newsletters, File 267 (1996 - present)

Set#	Query
L1	RANDOM$2 ADJ3 SLECT$4
L2	historical adj interval
L3	(RANDOM$2 ADJ3 SeLECT$4) same (historical$2 adj interval$2)

L4	(RANDOM$2 with SeLECT$4) and (historical$2 with interval$2) and ((financ$4 portfolio) with (analysis value valuation))

16/9/1 (Item 1 from file: 15)
02013639 52769245
Should options traders rely on stochastic volatility option pricing models?

Moore, Gary S; Patel, Rakesh
Derivatives Quarterly v6n3 pp: 23-30
Spring 2000
ISSN: 1081-3268 Journal Code: DRVQ
Document Type: Periodical; Feature Language: English Record Type: Fulltext Length: 8 Pages
Special Feature: Formula Table
Word Count: 4387

Abstract:

Recent research has raised the possibility that the famous volatility smile may be explained by models that correct for negative skewness and excessive kurtosis. An empirical test shows that negative skewness and excess kurtosis is not a necessary and sufficient condition for the presence of the usual volatility smile in stock options valued by the Black-Scholes option pricing model. The magnitude of the volatility smile found in this data is puzzling. Researchers and option traders would be well-advised to continue to look at other avenues of inquiry in their search for an explanations of the "smile" phenomenon. Option traders using these stochastic volatility models need to exercise due caution.

Text:

The empirically documented bias found in the Black-Scholes option pricing model [1973) known as the volatility smile (Rubinstein [1985)) has engendered a great deal of option pricing research. The volatility smile is a systematic pattern in the implied volatilities extracted from the Black-Scholes model. These are derived from an option on a single asset across different strike prices. Although somewhat time-dependent (Rubinstein [1985]), the most recent studies (Bakshi, Cao, and Chen [1997]) have typically shown that in-the-money options have the highest implied variance, followed by at-the-money options, and, finally, out-of-the-money options. It has been suggested that the volatility smile can be explained by incorporating additional elements of the stock return distributions in the option pricing models (Bates [1991, 1996]). It is the purpose of this article to raise a degree of skepticism with regard to these explanations.

Some authors have dubbed the pattern of the volatility smile a "smirk," because of recent changes in the pattern of the "smile." The pattern of the implied volatility derived from the Black-Scholes model found in the 1970s and early 1980s curled up at both ends. Bates [1996], among others, suggests that in the period following the 1987 crash, investors were more aware of the potential for downside movement of stocks. Consequently, the "smile" in the post-1987 period is more monotonic downward-sloping than in the past. In any event, empirical studies have confirmed a change in the basic shape of the "smile."
Bakshi, Cao, and Chen [1997] state,

As the smile evidence is indicative of negatively-skewed implicit return distributions with excess kurtosis, a better model must be based on a distributional assumption that allows for negative skewness and excess kurtosis.

Backus et al. [1998] find that excess kurtosis may impart a downward bias to BS implied volatilities. In their test of the BlackScholes model with the S&P 500 index, Bakshi, Cao, and Chen [1997] suggest their evidence is consistent with a model where the implied skewness has been negative and the kurtosis has been in excess of that found in normal distribution. An empirical test with these models by Bakshi, Cao, and Chen [1997] clearly indicates that implicit stock return distributions are negatively skewed with higher kurtosis than allowable in the Black-Scholes lognormal distribution. Furthermore, they point out that, in examining the annualized S&P 500, the negative skewness and high kurtosis are in contrast with the skewness (of zero) and kurtosis (of three) allowed by the lognormal distribution in the Black-Scholes model. Numerous other authors have developed models incorporating elements of stochastic volatility (e.g., Heston [1993], Hull and White [1987], and Bakshi, Cao, and Chen [1997]).

The implication of this research, that the failure of the Black-Scholes model to incorporate any degree of skewness and excess kurtosis of the stock market into the model may help explain the existence of the most commonly observed volatility smile, should be an important issue in option pricing. If true, the volatility smile would be much less of a mystery. Proof of the proposition would add strength to the notion that market traders should depend on option pricing models, which incorporate stochastic volatilities. Option traders could confidently adjust the model for these empirical facts and trade with an additional degree of confidence.

On the other hand, evidence against this proposition would cause traders using stochastic volatility models to exercise due caution, particularly during periods of market turbulence.

Of course, this raises the issue of whether the commonly observed volatility smile would still exist in a stock return distribution without negative skewness and excess kurtosis. Would the volatility smile be possible in a stock return distribution that is positively skewed and does not have high excess kurtosis? If so, then the volatility smile would remain a mystery. Moreover, in such a case, option research would be stimulated to go into other avenues.

It is not our contention that the degree of skewness and kurtosis is not an important input in asset or option pricing. We believe that it is. We are mainly interested in the contention that the usual pattern of the smile can be linked to a particular pattern in the skewness and kurtosis. We are not asserting that models that incorporate skewness and kurtosis are not more accurate than those that do not. These models should be more accurate, because they contain more degrees of freedom, and because it is well-known that stock return distributions are likely to be skewed.

A number of explanations have commonly been given for the bias in the Black-Scholes results. The most important is the problem of dividends. The original Black-Scholes [1973] option pricing model assumed that the stock upon which the option was written paid no dividends. It can be shown that a stock with no dividends should never be exercised early Black [1975] suggests an approximate procedure to adjust for the possibility of early exercise. Roll [1977], Geske [1979], and Whaley [1981] suggest a more exact procedure. In tests of the new dividend adjustment models, Geske and Roll [1983], as well as Sterk [1983], suggest that the effect of dividends can explain the bias in Black-Scholes [1973] option pricing. Noting that the

original model (Black-Scholes) did not include dividends, Sterk [1983] indicates that adjusting for dividends has a significant effect on reducing the bias of the model. However, the commonly observed volatility smile is still observed even after the BlackScholes model is adjusted for dividends (Bakshi, Cao, and Chen [1997]). Consequently the lack of dividend adjustments is an inadequate explanation for the commonly observed volatility smile.
Jarrow and Rudd [1982] suggest that the critical determinant of the price of a European stock option is the terminal stock distribution pattern. The classical BlackScholes [1973] model assumes that the distribution is lognormal. Jarrow and Rudd demonstrate four ways in which a true distribution can be different from the lognormal distribution while still having the same mean and standard deviation. They are: 1) both tails are thinner, 2) the left tail is fatter, while the right tail is thinner, 3) the left tail is thinner, while the right tail is fatter, and 4) both tails are fatter.

Jarrow and Rudd suggest that the thinner right and left tail model would cause the Black-Scholes model to overprice out-of-the-money and in-the-money calls and puts. A fatter left tail and a thinner right tail would cause the Black-Scholes model to overprice out-of-the-money calls and in-the-money puts, while causing underpricing for out-of the-money puts and in-the-money calls. A thinner left tail and a fatter right tail model would cause the Black-Scholes model to overprice out-of-the-money puts and in-the-money calls, and underpricing for in-themoney puts and out-of-the-money calls. When both tails are fatter, Black-Scholes would underprice out-of-themoney and in-the-money calls and puts. Jarrow and Rudd's article stimulated the development of option pricing models that could take some of these distributioninduced abnormalities into account.

Hull and White [1987] show that similar results to Jarrow and Rudd's could be obtained only when one takes into account the changes in volatility. They conclude that when the change in volatility is not correlated with the stock price, the Black-Scholes model would overprice options that are at the money or close to the money and underprice options that are deep in or deep out of the money This is because, as a stock's price increases, volatility decreases. As a result, it is unlikely that high stock prices will be gained. If the stock's price and the volatility are positively correlated, the Black-Scholes model tends to underestimate the price for out-of-the-money call options and overestimate the price for out-of the-money puts. In this case, because volatility increases as a stock's price increases, high stock prices are more probable than under geometric Brownian motion. As the stock's price decreases, volatility decreases, leading to the conclusion that a low stock price is less likely than under geometric Brownian motion.

The explanation of biases due to alternative terminal stock price distributions or the relationships between the volatility and the stock price remain interesting areas for future research. Any systematic bias could be the result of either the stochastic volatility or the alternative terminal stock price distributions, or a combination of the two.

Measurement error due to non-synchronous trading and reporting of options and stocks was suggested by Rubinstein to be a potential factor in explaining the biases in the Black-Scholes option model. Rubinstein [1985] indicates that the shape and magnitude of the volatility smile is the same in the post-October 1977 period, but is affected by dividends over the August 23, 1976-October 21, 1977, period when the volatility smile reversed its usual pattern. However, even using a vast amount of nearly synchronous data, Rubinstein [1985] and Bakshi, Cao, and Chen (1997] have not found any results that substantially differed from those found using less synchronous data. The noise caused by non-synchronous data has not been shown to be

systematic, and studies that eliminate the problem still show the presence of the usual significant volatility smiles. Therefore, nonsynchronous trading does not seem to be an adequate explanation for the volatility smile.

Transaction costs have been advanced by a number of authors, including Rubinstein [1985], as a possible explanation for the systematic bias, but even Rubinstein admits this explanation fails to explain the reversal of the implied volatility pattern he found in his seminal 1985 empirical piece. The broker's fees for trading in options did not change between the two periods of time that Rubinstein examined. Thus, transaction costs do not seem to be a likely candidate to explain the reversal of the volatility smile across periods. This also raises doubt as to whether transactions costs are an adequate explanation in any period.

Current thought remains that the bias may be related to a more complicated distribution pattern of the underlying stocks that the Black-Scholes model fails to incorporate. The Black-Scholes model assumes lognormal distributions for underlying stock returns, which in turn implies zero skewness and a kurtosis equal to three. The popularity of this hypothesis is shown by the number of models that adjust for these distributional assumptions. The finding of negatively implied skewness and excessive kurtosis and models that adjust for these facts has some following in the literature. Of course, this assumes that the underlying stocks have negative skewness and excessive kurtosis. If the underlying distribution does not possess these properties, any increased power of these models is due to their increased parameterization and not to clear theoretical advances.

Consequently, we examine a stock that is likely to have a positive skewness due to its great growth potential. If negative skewness and excess kurtosis is a necessary and sufficient condition for the presence of a volatility smile, then the smile should not exist in such a stock. If the explanation of these researchers is correct, we might expect to see a reversal of the usual volatility pattern in the face of positive skewness.

BILL GATES'S SMILE

Microsoft Corporation is an interesting subject for financial research. Not only is it unusual because of its significant growth and famous CEO, but its stock also has a very desirable property from the point of option pricing research. The stock has historically paid no dividends. Consequently, any volatility smile observed when applying the Black-Scholes option pricing to Microsoft cannot be attributable to a dividend effect. This is important because the mispricing of the Black-Scholes model has been explained by some researchers as related to a dividend effect (Sterk [1983]). When applying the model to Microsoft stock, one does not have to worry whether the dividend correction that has been added to the basic model is itself a source of misspecification.

Additionally, Microsoft options are very dense, meaning that they are highly liquid and traded quite frequently Indeed, in terms of trading volume, Microsoft options are among the highest on the exchange. This tends to decrease the significance of any non-synchronicity between the trading of the stock and the trading of the option. Closing trades at any given strike price usually occur very close to the closing bell.

Microsoft is often called a "very growth" stock, which means that it may have interesting distributional properties. Its beta is estimated by ValueLine to be 1.5, indicating that a high return is expected. As a consequence of the enormous upside potential in the software business, we suspect that the stock return distribution is positively skewed. This

result was empirically verified. Consequently, the distribution pattern represents an interesting contrast to a lower expected return stock or index of stocks.

DATA

We focus on Microsoft options, with data from two sources. Data provided by the Center for Research in Security Prices (CRSP) was used to examine the actual distributional properties of Microsoft returns over the January 1,1996-December 30, 1997, period. Closing option prices and the prices of the underlying stocks are taken from The Wall Street Journal. The underlying stock price is verified using the CRSP data.

We randomly selected a stratified sample of 485 options traded on Thursdays and Wednesdays during this period to avoid any potential turn-of-the-week effects. The sample was stratified so that a relatively equal number of in-the-money, at-the-money, and out-of themoney options were chosen. We applied screens similar to those used by Bakshi, Cao, and Chen [1997]. We eliminated any options that had less than eight days to maturity because these authors claim they are less liquid. We also eliminated any option trade that violated the rational boundary condition, as it is not possible to calculate an implied volatility on such an option. These screens resulted in the elimination of approximately 4% of our sample.

The riskless rate of return was computed from the average of the bid and ask yields reported in The Wall Street Journal for U.S. T-bills. The risk-free rate used is the continuously compounded yield having maturity closest to the expiration date of the option.
The implied volatility was computed using the Black-Scholes [1973] model. The basic formula behind the model is:

Since Microsoft options pay no dividends, no adjustment to the model is needed.

RESULTS

The average daily return for the two-year period starting on January 1996 is a very robust 0.0023286 (annualized 58% return). A potential important question is the difference between the historical volatility and the actual volatility; this has also been a potential problem in the Black-Scholes model. Unfortunately, there are three ways to compute the historical volatility, but they differ on how many volatility days are used. In the first method, the number of trading days in a year, 252, is used. So the historical daily volatility of Microsoft stock during the twoyear period from January 1996 is 0.019762.

Using the formula

Historical Volatility =

Daily Volatility (the number of trading days in a year) 1/2
We get

Annual Volatility = 0.0197262 (252) 112

Thus:

Annual Volatility = 0.3137

If we use the number of days in a year to compute historical volatility, the result is a bit different:

Historical Volatility = 0.019762 (365) 112 0.377524

The third way to calculate historical volatility involves using 258.5 days. French and Roll [1986] find that the variance of a typical stock over a Friday to Monday close is on average 1.11 times the variance over a weekday Similarly, a return over a three-day holiday weekend is on average 1.117 volatility days, and 1.107 if Monday is the holiday By this convention, there are 21.54 volatility days in the average month, and 258.5 volatility days per year. Substituting 258.5 days into the formula generates another slightly different number for historical volatility:

Historical Volatility = 0.019762 (258.5) 1/2 = 0.317732

Although using the number of trading days in a year seems the most reasonable method, the other two also have justifications. Nevertheless, how many days to use in calculating historical volatility is a debatable issue.

Actually, a number of issues remain controversial regarding the computation of historical volatility, such as: How many observations should one use? What should the time interval be between observations, and which price (opening, closing, high, or low) should we use? The selection of different approaches should be based on the purpose of the analysis, and the data availability of the users.

The lowest value of the historical volatility is 0.313722 over the entire period. The skewness of 0.31164 is very positive and the kurtosis is 1.739. Consequently, we conclude that the high growth rate of Microsoft has resulted in a return distribution that is non-negative and has no excess kurtosis. Consequently it would be interesting to know if the stock possesses the usual volatility smile seen in other recent studies.

We define the moneyness of the option as many authors have, as the stock price divided by the exercise price (see Choi and Wohar [1992]). Consequently, a moneyness measure of less than one means the option is out of the money, while a moneyness measure of more than one indicates the stock option is in the money

EXHIBIT 1

In examining Exhibit 1, we can see that the implied volatility pattern derived from the Black-Scholes model is the common one. The implied volatility of the out-of-the-money options is the smallest, followed by the at- or near-the-money options, and then by the inthe-money options with the largest implied variance. The implied volatilities seem relatively more consistent with historical volatilities calculated using 365 trading days, which are clearly larger than those calculated with 252 or 258.5 trading days. Although we can see the pattern by visual examination of the data, it was thought that a more rigorous test for the presence of a volatility smile was warranted.

A regression was run with the implied variability as the dependent variable and moneyness and time as the independent variables. The test about the residuals indicated a small problem that was corrected using the White (1980 robust estimator, which computes a consistent estimate of covariance matrix estimator. The White regression estimates are shown in Exhibit 2.

The coefficient on moneyness is positive, indicating the usual direct relationship. The t-test on the moneyness coefficient is approximately 9, indicating significance. The relationship between time and implied variance is inverse and statistically significant. The overall F-test is very significant. Thus, our data show a statistically significant documented "smile," with the usual pattern. The implied value for sigma for options written on a particular stock decline systematically as the exercise increases. A statistically significant inversely proportioned relationship

is found between implied volatility and time to maturity. The t-statistic on this variable is approximately -5, indicating a strongly significant relationship.

EXHIBIT 2

One interesting aspect of the regression is the coefficient on the predictor variable, moneyness. It seems very large. For purposes of comparison, a similar regression was run on a number of S&P 500 options over both similar and different periods. Based on a number of similar sample periods of similar-maturity options, the coefficient on moneyness on the S&P 500 options ranges from 0.17 to 0.52, with an average of 0.4. Our estimate of the slope of the moneyness regression is 0.75, which is almost double the S&P 500 index option coefficient.

This raises an interesting question. Is there a relationship between the expected return on a stock and the slope of the volatility smile? We have intriguing evidence that the answer might be yes. Although the magnitude of the moneyness coefficient is not of direct interest to the research questions developed here, the finding is quite interesting. However, given the confusing behavior of the historical shifts in the volatility "smile," we await future research on this issue.

These results are somewhat puzzling if negative skewness and excessive kurtosis explains the volatility smile. We expect the smile to reverse its usual pattern in the face of the positive skewness seen in the Microsoft stock. Our only conclusion is that negative skewness and excessive kurtosis are not necessary conditions for the presence of an implied volatility smile. This is a useful but limited finding.

DISCUSSION

Why does a positively skewed stock with low kurtosis still have such a strong volatility smile? Our research raises the question of whether the additional parameterization suggested by many authors is really worth the price. Is the volatility smile more a function of the time period than the characteristics of the expected terminal stock distribution? Bakshi, Cao, and Chen [1997] suggest that the use of stochastic volatility can be justified based on the superior out-of-sample predictive power of stochastic volatility models. Comparison of the stochastic volatility model with the Black-Scholes model seems to show that the overall pricing performance of stochastic volatility models is superior. Bakshi, Cao, and Chen find that allowing for negative jumps is useful insofar as it increases the skewness of the distribution of the stock, but that the jump process does not generate the level of skewness implied by the volatility smirk observed in the empirical option data. No model they tested adequately explained the actual empirical option data.

Clearly, the presence of the smile bias as well as the improvement in the out-of-sample predictive power found by these researchers shows that there is an omitted variable in the Black-Scholes model. The question is whether we have found the right omitted variable.

Sabbatini and Linton [1998] report superior performance of a GARCH option pricing model. A GARCH (1, 1) model is a good parameterization of the process. However, large out-of-sample pricing errors suggest that the results should be taken with caution.

Given our results, we suggest that the superior predictive power of stochastic volatility models be viewed with an equal amount of skepticism. It may be that the superior power of the stochastic model is more a function of its greater degrees of freedom than a superior economic

explanation of the underlying relationships.

CONCLUSION

According to recent research, the separation of the pattern of skewness and kurtosis of stock market returns from the idealized return distribution of the BlackScholes model has raised a possible explanation for the famous volatility smile. Suggested models that correct for negative skewness and excess kurtosis have been developed for the purposes of clarifying the known empirical biases in option pricing. It is our purpose to raise a degree of skepticism with regard to these explanations.

In this context, the presence of the commonly observed volatility smile in Microsoft is significant because the stock does not have negative skewness and excessive kurtosis during the periods tested. Indeed, the scope of the bias as measured by the slope of the coefficient on moneyness indicates an extremely powerful economic effect.

As evidenced by an empirical test using the BlackScholes model, negative skewness and excessive kurtosis are not a necessary and sufficient condition for the presence of an implied volatility smile in stock options. Because the stock pays no dividends, the volatility smile cannot be explained as in some past studies as a lack of proper dividend adjustment. Continued exploration into other areas of investigation would likely produce an alternative explanation for this phenomenon. Until then, the mystery continues.

Our findings have obvious implications for practitioners who currently use stochastic implied volatility models. Option traders using these models need to exercise due caution. Although the additional parameterization built into these stochastic volatility models may impart useful information to traders, reliance upon the models during times of market turbulence may be quite unwise.

REFERENCES

Backus, David, Steve Foresi, Li Kai, and Liuren Wu. "Accounting for Biases in Black-Scholes." Working paper, New York University, 1998.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. "Empirical Performance of Alternative Option Pricing Models." Journal of Finance, 52 (1997), pp. 2003-2049.

Bates, David. "The Crash of '87: Was it Expected? The Evidence from Options Markets."Journal of Finance, 46 (1991), pp. 1009-1044.

_. "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutschemark Options." Review of Financial Studies, 9 (1996), pp. 69-108.

Black, Fischer. "Fact and Fantasy in the Use of Options." Financial Analysts Journal, Vol. 31, No. 4 (1975), pp. 36-41. Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities."Journal ofPolitical Economy, 81 (1973), pp. 637-659.

. "The Valuation of Option Contracts and a Test of Market Efficiency." Journal of Finance, May 1972, pp. 399-417. Choi, Seungmook, and Mark E. Wohar. "Implied Volatility in Options Markets and Conditional Heteroscedasticity in Stock Markets." The Financial Review, Vol. 27, No. 4 (1992), pp. 503-530.

French, Kenneth, and Richard Roll. "Stock Return Variance: The Arrival of Information and the Reaction of Traders." Journal of Financial Economics,

Vol. 17 (1986), pp. 5-26.

Geske, Robert. "A Note on an Analytic Valuation Formula for Unprotected American Call Options on Stacks with Known Dividends. "Journal of Financial Economics, 7 (1979), pp. 375-380.

Geske, Robert, Richard Roll, and Kuldeep Shastri. "Over-theCounter Option Market Dividend Protection and 'Biases' in the Black-Scholes Model: A Note." Journal of Finance, Vol. 38, No. 4 (1983), pp. 1271-1277.

Heston, Steven. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." Review of Financial Studies, 6 (1993), pp. 327-343. Hull, John, and Alan White. "The Pricing of Options with Stochastic Volatilities." Journal of Finance, 42 (1987), pp. 281-300.

Jarrow, Robert, and Andrew Rudd. "Approximation Option Valuation for Arbitrary Stochastic Processes." Journal of Financial Economics, 10 (1982), pp. 347-369.

Roll, Richard. "An Analytic Formula for Unprotected American Call Options on Stocks with Known Dividends." journal of Financial Economics, 5 (1977), pp. 251-258.

Rubinstein, Mark. "Nonparametric Tests of Alternative Option Pricing Models Using all Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978." Journal of Finance, 40 (1985), pp. 455-480.

Sabbatini, Michael, and Oliver Linton. "A GARCH Model of the Implied Volatility of the Swiss Market Index from Option Prices." InternationalJournal of Forecasting, Vol. 14, No. 2 (1998), pp. 199-213.

Sterk, William Edward. "Option Pricing: Dividends and the Inand Out-of-the-Money Bias." Financial Management, Vol. 12, 4 (1983), pp. 47-53.

Whaley, Robert, and Robert Geske. "On the Valuation of American Call Options on Stacks with Known Dividends." Journal of **Financial** Economics, 9 (1981), pp. 207-215.

White, Halbert. "A Heteroskedasticity-Consistent Covariance Estimator and Direct Test for Heteroskedasticity." Econometrica, 48 (1980), pp. 817-838.

GARY S. MOORE is director of the College of Business Honor's Program in the department of finance at the University of Toledo in Ohio.

RAKESH PATEL is a visiting scholar at Washington University in Washington, D.C.

THIS IS THE FULL-TEXT.

Geographic Names: United States; US

Descriptors: Stock options; Volatility; Mathematical models; Studies
Classification Codes: 3400 (CN=Investment analysis & personal finance); 9130 (CN=Experimental/Theoretical); 9190 (CN=United States)
Print Media ID: 33269

16/9/2 (Item 2 from file: 15)

00956845 96-06238

An exploratory analysis of portfolio managers' probabilistic forecasts of stock prices

Muradoglu, Gulnur; Onkal, Dilek

Journal of Forecasting v13n7 pp: 565-578
Dec 1994
CODEN: JOFODV
ISSN: 0277-6693 Journal Code: JOF
Document Type: Journal article Language: English Length: 14 Pages
Special Feature: Charts Appendix Equations References
Word Count: 6320

Abstract:

An experiment examined the effects of forecast horizon on the performance of probability forecasters and the alleged existence of an inverse expertise effect, i.e., an inverse relationship between expertise and probabilistic forecasting performance. Portfolio managers are used as forecasters with substantive expertise. Performance of this "expert" group is compared to the performance of a "semi-expert" group composed of other banking professionals trained in portfolio management. It is found that while both groups attain their best discrimination performances in the 4-week forecast horizon, they show their worst calibration and skill performances in the 12-week forecast horizon. Also, while experts perform better in all performance measures for the one-week horizon, semi-experts achieve better calibration for the 4-week horizon. It is concluded that these results may signal the existence of an inverse effect that is contingent on the selected forecast horizon.

Text:

Forecasting accuracy of financial variables attracts considerable research attention, with predominantly conflicting findings. Liljeblom (1989) shows expirical evidence displaying the predictive power of financial analysts' forecasts of earnings per share for short-term forecast horizons in the Scandinavian securities market. DeBondt and Thaler (1990), on the other hand, conclude that the earnings forecasts provided by the security analysts are basically overreactive, i.e., when large earnings increases are forecasted actual earnings are lower than predictions and vice versa. Keane and Runkle (1990) use the ASA-NBER survey of economic forecasters and show that the price forecasts for GNP deflator by expert forecasters are rational demonstrating improved performance. Zarnovitz (1985) uses the same data set and rejects the rational expectations hypothesis for inflation forecasts. Summarizing the relevant work on financial forecasting, DeBondt (1991) concludes that 'finance should attempt to model the behavior of representative investors and the nature of their errors' (p. 90). Contradictory results obtained in these studies may be viewed as resulting from data revisions, biases in aggregating data, and the definitions of forecast error that are employed. Also, all the studies mentioned above focus on investigating the accuracy of point forecasts. (For discussions on enhancing point forecasts with prediction intervals and the associated complexities, see Klein, 1971; Granger and Newbold, 1986.) Probabilistic

forecasting offers an alternative approach to forecasting in financial settings. Whereas the accuracy evaluations of point forecasts are quite limited, the use of probabilistic forecasts enables a detailed analysis of the various performance characteristics, of forecasters.

Probabilistic forecasting tasks demand the assessment of subjective probabilities as articulations of the forecaster's degrees of belief in the occurrences of future outcomes. When viewed in this framework, probabilistic forecasts accord quantitative descriptions of forecasters' uncertainty (Murphy and Winkler, 1974). They also provide a channel for communicating this uncertainty to the users of the forecasts, who can, in turn, better 'interpret' the forecasts and make more informed decisions by assimilating the uncertainty associated with the forecasts. In addition to providing the information transmitted by point or categorical forecasts, probabilistic forecasts offer mechanisms that forecasters could utilize to convey their true judgments, hence reducing any biasing tendencies (Daan and Murphy, 1982). In short, it is argued that probabilistic forecasts are more useful than point or categorical forecasts, since they provide more detailed information to the users (Murphy and Winkler, 1992).

Within the realm of forecasting stock prices, subjective probability distributions were initially used by Bartos (1969) and Stael von Holstein (1972), with the common finding that the uniform distribution outperforms the distributions provided by various forecasters. Bartos (1969) replaced historical prices in the Markowitz (1959) mean-variance portfolio method by density functions and cumulative distribution functions assessed by security analysts, but could not delineate a better portfolio performance. Stael von Holstein (1972) investigated the effects of feedback on forecasting performance by using non-dichotomous distributions, but the forecasts could not outperform those based on past frequencies. Both studies concluded that further research on selection of appropriate forecast horizons is required.

More recently, Yates et al. (1991) reported experimental results on the overall inaccuracy of probabilistic forecasts of earnings per share and stock prices. Their work was especially important for their conjecture of the 'inverse expertise effect'. Yates et al. found an inverse relationship between expertise and forecast accuracy. Linking their findings to Hammond's (1966) work on probabilistic functionalism, the authors suggested that the inverse expertise effect is a byproduct of the experts' cue utilizations. That is, it is asserted that, in their decision-making processes, experts use a significantly larger number of cues than non-experts. This in turn, is believed to lead them to consolidate irrelevant cues into their judgment processes, thereby decreasing the accuracy of experts' forecasts.

The alleged inverse expertise effect was further found in a similar experiment utilizing only price forecasts (Onkal and Muradoglu, 1994). Although both studies used students as forecasters, Yates et al. (1991) specified graduate students as experts, while Onkal and Muradoglu (1994) defined experts as students who have previously made stock investment decisions. Also, while Yates et al. employed a three-month forecast horizon in a developed securities market, Onkal and Muradoglu used a one-week horizon in an emerging market. Yates et al. used quarterly forecasts assuming that the two-week forecast horizon employed by Stael von Holstein (1972) was shorter than customary for professional forecasters. Bartos (1969) had used one-month, three-month, and six-month forecast horizons, but he found it difficult to make generalizations due to sample size restrictions. In order to reduce task complexity as well as to mimic the current practices of financial media, Onkal and Muradoglu (1994) used a forecast horizon of one week.

Considering the utilization of different time horizons in the studies reviewed above, it may be claimed that the choice of forecast horizon is a critical variable that may have a powerful impact on experimental results. This study aims to investigate the effects of forecast horizon on the performance of probability forecasters. In doing so, it seeks to examine the alleged existence of the inverse expertise effect by using probabilistic forecasts provided for different time horizons by portfolio managers. Portfolio managers are used as forecasters with substantive expertise. Performance of this 'expert' group is compared to the performance of a 'semi-expert' group entailing other banking professionals trained in portfolio management. To construct these comparisons, the experimental framework of Yates et al. (1991) is adapted to an inherently less complicated decision environment, and only forecasts of stock prices are used.

One of the primary contributions of this research is that it employs professional fund managers as forecasters with substantive expertise. Second, various performance aspects of these expert and corresponding semi-expert groups are compared with respect to differing forecast horizons. Third, the particular setting used (i.e., a developing economy setting with an inefficient emerging security market) avoids any misinterpretations of forecast accuracy through assumptions regarding market efficiency. Whereas previous research has attempted to explain the relatively poor forecasting performance of experts via market behaviour (i.e., conditions of market efficiency), our research aims to analyse forecasting performance via experts' and semi-experts' expressions of uncertainty.

MARKET EFFICIENCY AND PORTRAYAL OF THE EXPERIMENTAL SETTING

A market is defined to be efficient with respect to an information set if it is not possible to make abnormal profits by adopting trading strategies using this information set (Jensen, 1978). The semi-strong form of market efficiency uses publicly available information in addition to past prices as the relevant information set. Therefore, if a market is efficient in the semi-strong form, expert-managed portfolios, for example, are not expected to beat the market since the information set is equally accessible to all (Jensen, 1968). In this case, 'the basic question most asked is--are price changes forecastable?' (Granger, 1992, p. 3). There is contradictory evidence from finance literature regarding the performance of expert-managed funds (Elton et al., 1991; Ippolito, 1989). When expert-managed funds do not beat the market, this is typically attributed to the efficiency of the market, i.e., an explanation is provided via market behaviour rather than forecaster performance.

It may be argued that the existence of an inverse expertise effect can be used as a behavioural explanation to the inconclusive results of semi-strong form efficiency tests using professionally managed portfolios. The implication is that the setting used in the current study (i.e., a developing economy with an inefficient emerging security market (Sengul and Onkal, 1992; Unal, 1992)) prohibits any misinterpretations of forecast accuracy via market-efficiency assumptions.

In particular, the financial markets in Turkey were strictly regulated until 1980, when the IMF and the World Bank supported the introduction of a liberalization package. In 1982 the required legal framework and regulatory agencies for the stock market were established. Istanbul Securities Exchange, the only stock exchange in Turkey, started operations in 1986. Employees of the stock exchange could hold stock portfolios without notification until 1988, and there was no legislation against insider trading until 1990. When this study was conducted (i.e. February 1992), 143 stocks were traded at the Istanbul Securities Exchange, and the average daily volume of trade was US$55 million. There was a total of 162 intermediaries and brokerage houses, 60 of which were affiliated with

companies traded in the exchange.

PROCEDURES

All the subjects in the study were professionals in the stock market. No monetary or nonmonetary bonuses were offered to the participants. The current study was depicted as one that suggests an alternative method of presenting stock price forecasts. Concerns on the evaluation of forecasting performance and reliability of forecasts were discussed as primary issues involving portfolio management. Probabilistic forecasts were presented as pertinent channels of communication between forecasters and users of forecasts. It was argued that if forecasters could employ probabilistic forecasts as quantitative descriptions of their uncertainty, it may be possible to reduce biases such as overforecasting.

Participants of the study were reached at two different locations at the same date. The first group, called 'semi-experts', was composed of two internal auditors and eight managers who had completed a company-paid 40-hour training programme on portfolio management. The internal auditors were trained in portfolio management since they were going to be specializing in the auditing of bank-affiliated brokerage houses. All the managers in this group were employed in banks. They attended the training programme because they were expected to work for bank-affiliated brokerage houses upon completion of the programme.

The second group, called 'experts', was composed of seven portfolio managers working for a bank-affiliated brokerage house. All the experts had licences as brokers and their job descriptions included managing investment funds and giving investment advice to customers with investments above US$50,000.

The task was defined as the preparation of probabilistic forecasts of the closing stock prices for 34 companies listed on the Istanbul Stock Exchange. Companies had been selected on the basis of their volume of trade on the preceding 52-week period. This particular selection was made to minimize task complexity as the stocks with relatively high volumes of trade could be followed with minimum effort by all experts and semi-experts.

Participants were requested to make probabilistic forecasts regarding the price changes for the 34 stocks for forecast horizons of one week, two weeks, four weeks, and twelve weeks. In particular, the forecasts were to be made regarding the percentage change between the previous Friday's closing stock price and (1) the closing stock price that will be realized on the Friday due in one week, (2) the closing stock price that will be realized on the Friday due in two weeks, (3) the closing stock price that will be realized on the Friday due in four weeks, and (4) the closing stock price that will be realized on the Friday due in 12 weeks.

Subjects were requested to provide these forecasts in the form of subjective probabilities conveying their degrees of belief in the actual price change falling into the designated percentage change categories. Specifically, they were asked to complete the following response form for each stock for each forecast horizon:

(Response form omitted)

The range of stock price changes in the response form were formulated by considering the average weekly, bi-weekly, monthly, and quarterly price changes of the composite stock index during the previous 52-week period. For the past 52 weeks, weekly price changes were 3% on average with the maximum increase being 8% and the maximum decrease 5%. During the same

period, the average quarterly price change was 13% with the maximum increase being 65% and the maximum decrease 21%. The first 5% increase range (Interval 5) contained the average weekly increase during the previous year, the second (Interval 6) contained the maximum average weekly change, the third (Interval 7) contained the average quarterly change, and the fourth range (Interval 8) was designed for stocks whose quarterly volatility could be higher than average. Intervals 1 to 4 were designed to be symmetric to intervals 5 to 8 for cognitive purposes. Subjects were not given different intervals for different forecasting periods in order not to convey and confound the personal opinions of the researchers regarding the expected changes for differing forecast horizons.

In the beginning of the experiment, concepts related to subjective probabilities and probabilistic forecasting tasks were discussed with detailed examples. Design and goals of the study were depicted. Participants were not given any information about the results of similar studies on inverse-expertise effect (e.g., Onkal and Muradoglu, 1994; Yates et a)., 1991) so that their motivation would not be affected.

Subjects were informed that certain scores of probabilistic forecasting performance would be computed from their individual forecasts. They were advised that, due to the computational characteristics of these scores (i.e., proper scoring rules), participants could earn their best potential scores by expressing their true opinions and thereby avoiding hedging or bluffing. Subjects were told that they would be notified about their forecasting performance on a personal basis, and no information about their direct or implied individual performances would be provided to their managers or co-workers.

Each subject was provided with background folders for each of the 34 companies, delineating company name, industry, net profits as of the end of the third quarter of 1991, earnings per share, and price--earnings ratios as of the last day of the preceding week. Also provided were the weekly closing stock prices (i.e., the closing stock prices for each Friday) of the preceding 52 weeks in graphical form. Weekly closing stock prices for the last 3 months (12 weeks) were also presented in tabular form. Subjects were permitted to utilize any source of information other than the remaining participants of the study.

In order to duplicate real forecasting settings, subjects were allowed to take the background folders home. They were given the background folders and response sheets on Friday afternoon (after the session has closed and the closing prices were known) and were requested to submit the completed response sheets by Monday 9 a.m. (before the opening of the next session at the stock exchange). Participants were to render their forecasts using the response forms illustrated previously.

RESULTS

Forecasting performances of experts and semi-experts for the various forecast horizons are investigated using the following measures (see the Appendix for a detailed discussion of these performance indices):

(1) Mean probability score (abbreviation omitted)--presents a measure of overall forecasting accuracy;

(2) Calibration--provides an index of the forecaster's ability to match the probabilistic forecasts with the relative frequencies of occurrence;

(3) Mean slope--gives a measure of the forecaster's ability to discriminate between instances when the realized price change will or will not fall into the specified intervals;

(4) Scatter--conveys a measure of excessive forecast dispersion;

(5) Forecast profile variance--compares how close the forecaster's probability profile is to a flat profile that shows no variability across intervals; and

(6) Skill--presents an index of the overall effect of those accuracy components that are under the forecaster's control.
The analyses utilizing these measures are conducted at three levels. First, performances across different forecast horizons are tracked within each group by using median tests. Next, for each of the forecast horizons, comparisons of experts and semi-experts are made using Mann-Whitney U-tests for each of the performance measures. Finally, scores of experts and semi-experts are compared to those that would be obtained by the uniform, historical, and base-rate forecasters (i.e., the three comparison standards for forecasters as suggested by Yates et al., 1991). The uniform forecaster presents a comparison basis for forecasters, since it makes no discriminations among intervals, and hence, assigns equal probabilities to all the intervals (i.e., probability of 1/8 to each of the eight **intervals** for each stock). Second standard of comparison is given by a **historical** forecaster, who provides forecasts identical to the historical relative frequencies. Given the volatility of the stock market under consideration, the historical forecaster's probability forecasts are set equal to the relative frequencies realized in the previous week. The third comparison standard is provided by the base-rate forecaster. This forecaster is analogous to a clairvoyant who can perfectly foresee the relative frequencies (i.e., base rates) with which the price changes will occur for that forecast horizon.

Table I presents the medians for the six performance measures for experts and semi-experts, along with the scores that would be obtained by the uniform, historical, and base-rate forecasters, for the four forecast horizons of interest. (Table I omitted) As can be observed from this table, experts clearly outperform semi-experts for the one-week forecast horizon. In particular, experts attain better scores in (abbreviation omitted) ($p = 0.026$), calibration ($p = 0.016$), mean slope ($p = 0.016$), scatter ($p = 0.049$), and skill ($p = 0.026$). Forecast profile variances of experts and semi-experts for this forecast horizon are very similar.

Experts and semi-experts also show very similar performances in all measures for the two-and 12-week forecast horizons. It is in the four-week forecast horizon that an inverse expertise effect with respect to calibration is observed. Calibration performance of semi-experts is found to be significantly better than that of the experts for the four-week horizon ($p = 0.012$). This indicates that semi-experts' probabilistic forecasts corresponded more closely with the realized relative frequencies for that particular forecast horizon.

In fact, the four-week forecast horizon appears to be the optimal time specification for the forecasters' discrimination performances, as indexed by their mean slope scores. This is true regardless of the forecasters' levels of expertise. For the expert group, mean slope scores obtained in the four-week horizon are better than the scores obtained in the one-, two-, and 12-week horizons (all $p = 0.015$). Also, calibration and skill performances of our expert participants are found to be at their worst level for the 12-week horizon. Calibration and skill performances for the 12-week period is worse than the corresponding performances in one, two, and four weeks (all $p = 0.0003$ for calibration; and all $p = 0.015$ for skill). As a result of these findings, forecasting accuracy (as indexed by the mean probability score) for the four-week horizon is found to be better than that of the 12-week horizon ($p = 0.015$).

It is interesting to note that the identical conclusions prevail for our semi-experts. Mean slope for the four-week horizon is found to be superior for semi-experts than that for the one-week ($p = 0.0006$), two-week ($p = 0.089$), and 12-week ($p = 0.012$) horizons. Calibration and skill scores are again found to be the worst for the 12-week horizon. Specifically, the calibration performance in 12 weeks is worse than the performance displayed in one week ($p = 0.0006$), two weeks ($p = 0.0000$), and four weeks ($p = 0.0000$). Likewise, this group's skill performance for the 12-week horizon is surpassed by its performance in the one-week ($p = 0.089$), two-week ($p = 0.012$), and four-week ($p = 0.012$) horizons. As a result of these performances, the mean probability scores for semi-experts for the four-week period are better than the 12-week period ($p = 0.089$).

We also accumulate evidence for an inverse expertise effect that is contingent on the forecast horizon, when we compare the performances of experts and semi-experts with those of the uniform, historical, and base-rate forecasters. As can be observed from Table II, the overall performance of 43% of experts are better than the uniform forecaster only for the one-week forecast horizon, while semi-experts demonstrate improved overall performance in comparison to the uniform forecaster for longer forecast horizons. (Table II omitted) Also, 30% of semi-experts show an overall accuracy level better than that of the historical forecaster for the four-week forecast horizon. This may be viewed as signalling that the use of subjective probabilities instead of the customary use of historical data could result in higher profit opportunities.

Regarding the calibration performance, it can be deduced from Table II that experts' ability to assign probabilities matching the actual relative frequencies of future outcomes deteriorate for longer forecast horizons. In contrast, the semi-experts seem to show improvement for longer horizons, with the most pronounced difference surfacing in the four-week forecast horizon.

Mean slope scores indicate that the semi-experts' ability to discriminate between instances when the actual price change will and will not fall in the specified intervals ameliorates more than that of the experts for the four-week forecast horizon. This finding persists for all benchmarks, including the uniform, historical, and base-rate forecasters.

It can be seen from Table II that, while 43% of the experts show better skill than uniform forecaster for the one-week forecast horizon only, semi-experts attain better skill scores as the forecast horizon is extended, with better performances than the historical forecaster for the four-week horizon.

DISCUSSION AND CONCLUSION

This research was focused on analysing the probabilistic forecasts of stock prices given by portfolio managers and other banking professionals in an emerging securities market setting. Goals of the study included (1) exploring the potential effects of varying forecast horizons on different dimensions of forecasting accuracy, and (2) investigating the alleged existence of the so-called inverse expertise effect. A significant contribution of this work involved the critical finding that the previous assertions regarding the so-called inverse expertise effect (Stael von Holstein, 1972; Yates et al., 1991; Onkal and Muradoglu, 1994) could be expanded to incorporate a dependence on the selected forecast horizon. As amplified above, the portfolio managers and the other banking professionals participating in the current study show distinctly different performances for the varying forecast horizons of interest.

Our findings suggest that the forecast horizon yielding the best

discrimination performance for both experts and semi-experts is four weeks under the current emerging securities market setting. All the participants show superior performance (as displayed by the mean slope scores) in discriminating between occasions when the actual price change will or will not fall into the specified intervals for the four-week horizon in comparison to the one-, two-, and 12-week horizons. Also, regardless of the level of expertise, calibration and skill scores for all participants reach their least desirable values for the 12-week forecast horizon. These results may be viewed as implying that the forecasters' ability to assign appropriate probabilities to actual outcomes as well as their overall forecasting skill may be better for shorter forecast horizons. It may also be that the poorer performance observed for 12 weeks may relate to caution or 'mean reversion' in forecasting which could be vindicated over a longer time series.

Comparisons of the portfolio managers' performances with those of other banking professionals suggest that the performance of experts becomes worse than that of semi-experts as the forecast horizon is extended. For the one-week forecast horizon, the performance of experts are significantly better than that of semi-experts. For the two-week forecast horizon there is no difference between the forecast performance of portfolio managers and semi-experts. However, for the four-week forecast horizon, which appears to be the optimal period for discrimination performances, it is observed that the calibration scores of semi-experts are better than that of experts; i.e., banking professionals' skills in assigning probabilities that match the realized relative frequencies are better than that of portfolio managers.

Previous research has adopted two main explanations for the inverse expertise effect. The first explanation stems from the assertion that experts use more abstract representations of the problem situation, hence causing novices to outperform the experts (Adelson, 1984). The second suggests that since experts use richer representations (Murphy and Wright, 1984), their use of additional cues make the judgment task more difficult for them, thus distorting the accuracy of their judgments (Yates et al., 1991). Current results may be viewed as implying that the representations used by portfolio managers, along with their cue utilizations, may change with the forecast horizons. It may be argued that, since portfolio managers predominantly make short-term forecasts in emerging markets with higher volatilities, shorter forecast horizons provide a better fit to these experts' natural environments. This leads to a superior forecasting performance for the one-week forecast horizon. However, as the forecast horizon becomes longer, there emerges a discrepancy with the experts' natural domain. Portfolio managers may be viewed as responding to this discrepancy by showing a deteriorated ' calibration performance, thus further confirming the arguments of the ecological approach that probability assessors are well calibrated to their natural environments (Gigerenzer et al., 1991; Juslin, 1993). Relatedly, portfolio managers are not found to be successful in dispersing their uncertainty through intervals for longer forecast horizons. Instead, they choose to concentrate on a few intervals, and their calibration scores are distorted.

Financial forecasts are generally reported in the form of point or categorical forecasts that do not disclose how firmly the forecasters believe in their expectations. Griffin and Tversky (1992) have asserted that experts could be expected to become more overconfident than novices in unpredictable domains like the stock market, since they are more likely to give unwarranted credibility to their fallible expert knowledge. Portfolio managers' potential miscalibration and the resulting overconfidence may mislead the users of financial forecasts in several ways. First, investment advice supplied by these experts may result in investors forming less diversified portfolios. Second, the long-run performance of portfolios managed by these experts may not be different from **randomly selected** others or benchmarks utilizing some normal return

criteria. Inefficiency of the market in strong form may in fact be a signal of the expert's behavior.

Previous research has repeatedly emphasized the critical role played by judgment in economic and financial forecasting (Wright and Ayton, 1987; Batchelor and Dua, 1990; McNees, 1990; Turner, 1990; Wolfe and Flores, 1990; Bunn and Wright, 1991; Flores, Olson and Wolfe, 1992; Donihue, 1993). The use of probabilistic forecasts provides a means for consolidating the uncertainty inherent in the stock market into financial forecasts. The users of financial information are already accustomed to probability distributions, which may in turn be effectively utilized to convey the forecasters' degrees of uncertainty. The providers of financial forecasts can employ probabilities as mechanisms for assessing and evaluating their uncertainty, hence furnishing more detailed information to the users. In short, it is our belief that probabilistic forecasting constitutes an important channel of communication between the providers and the users of forecasts, and that it demands efficient utilization in financial settings.

We would expect future work in this area to concentrate on the following research avenues. First, the differential performances observed for the forecast horizons in this study may in part be a product of the specific conditions of the financial environment. Therefore, comparisons of portfolio managers' forecasts in different settings need to be made by controlling for the volatility of the stock markets and the efficiency conditions. Second, similar experiments need to be conducted in bull and bear markets to control for the effects of cyclical movements in stock markets. Finally, feedback studies with portfolio managers are required to examine whether the forecast-horizon-dependent inverse expertise effect found in the current study will disappear with relevant training and feedback.

Following the framework of O'Connor (1989), performance in probabilistic forecasting tasks may be viewed as a function of the participants' familiarity with the topic, experience with subjective probabilities, availability of feedback, and the environmental context. Accordingly, delineating the implications of these factors constitutes a critical threshold for effective applications of judgmental forecasting in financial settings.

APPENDIX: PERFORMANCE MEASURES USED

1. Probability score for multiple events

The forecast vector given by a subject for each stock is defined as f = (f sub 1 , f sub 2 , f sub 3 , f sub 4 , f sub 5 , f sub 6 , sub 7 , f sub 8), where f sub k denotes the probabilistic forecast that the stock's price change will fall into interval k, k = 1,2,...,8. The outcome index vector is defined as d = (d sub 1 , d sub 2 , d sub 3 , d sub 4 , d sub 5 , d sub 6 , d sub 7 , d sub 8), where d sub k assumes the value of 1 if the realized price change falls within interval k and the value of 0 if it does not. The scalar product of the difference between the forecast vector and the outcome index vector gives the probability score for multiple events (PSM). That is:

PSM = (f - d)(f - d) sup T = Sigma (f sub k - d sub k) sup 2

As defined above, the range of PSM is [0, 2]. The lower this value, the better is the forecaster's accuracy with respect to the particular stock in question. Accordingly, a forecaster's overall accuracy level can be indexed by taking the mean of the probability scores (abbreviation omitted) over a specified number of forecasting instances (i.e., over a given number of stocks). Components emanating from the Yates decomposition of (abbreviation

omitted) (as explained below) are employed a indices of various performance attributes of forecasters (Yates, 1988).

2. CALIBRATION

Calibration provides information about the forecaster's ability to assign appropriate probabilities to outcomes. A forecaster is well calibrated if for all predicted outcomes assigned a given probability, the proportion of those outcomes that occur (i.e., proportion correct) is equal to the probability. For example, suppose that, over 100 predictions, a forecaster assesses a probability of 0.3 that the given stock's price will increase by more than 15%. This forecaster's 0.3 assessments are well calibrated if an increase of more than 15% is actually observed on 30 of the 100 predictions. If the forecaster's other probability forecasts similarly match event frequencies, the forecaster is said to be well calibrated. Accordingly, a calibration score can be computed as a function of (characters omitted), representing the mean probability forecast for interval k, and (characters omitted), representing the proportion correct for interval k:

(Equation omitted)
Lower scores refer to better calibration performance.

3. MEAN SLOPE

Mean slope provides an indication of the forecaster's ability to discriminate between occasions when the actual price change will and will not fall into the specified intervals. Higher values of the mean slope imply finer discrimination. The computation of mean slope is as follows:

(Equation omitted)

where (characters omitted) is the mean of probability forecasts for a price change falling into interval k computed over all the cases where the realized price change actually fell into interval k. Similarly, (characters omitted) is the mean of probability forecasts for a price change falling into interval k computed over all the times when the realized price change did not fall into the specified interval.

4. SCATTER

Scatter is that part of the overall forecast variance that is not attributable to the forecaster's ability to discriminate between occasions when the actual price change will or will not fall into the specified intervals. The best value of the scatter would be zero, since it signals excessive variance due basically to the forecaster's reaction to nonpredictive environmental cues. A scatter index could be computed as follows:

Scatter = Sigma Scatter sub k = Sigma (1/N) [(N sub 1k * Var(f sub 1k)) + (N sub 0k * Var(f sub 0k))]

where Var(f sub 1k) is the conditional variance of the N sub 1k forecasts given for a price change falling into interval k when it actually occurred. Similarly, Var(f sub 0k) is the conditional variance of the N sub 0k forecasts given for a price change falling into interval k when it did not. Of course, N = N sub 1k + N sub 0k .

5. FORECAST PROFILE VARIANCE

Forecast profile variance measures the discrepancy between a forecaster's set of probabilities and a uniform set of probabilities. Hence, the

forecast profile variance compares the forecaster's probability profile with a fat profile that shows no variability across intervals. An index of the forecast profile variance (for our study employing eight intervals) could be computed as follows:

(Equation omitted)
This index provides a measure of how different the forecaster's probabilities are from the nondiscriminating probabilities of the uniform forecaster.

6. SKILL

The overall effect of those (abbreviation omitted) components under the forecaster's control can be indexed via a skill score, which can be computed as follows:

(Equation omitted)

where Var(d sub k) is the variance of the outcome index d sub k for interval k. Since d sub k is determined by what happens in the forecasting environment (i.e. the realized price change), (equation omitted) indexes an uncontrolable element of (abbreviation omitted). Subtracting this 'base-rate' component from (abbreviation omitted), we have the overall effect of those components that are under the forecaster's control. Hence, lower skill scores imply better overall forecasting quality as displayed by the probability forecasts .

ACKNOWLEDGMENTS

We would like to thank Derek Bunn and two anonymous reviewers for their helpful comments and valuable suggestions.

REFERENCES

Adelson, B., 'When novices surpass experts: The difficulty of a task may increase by expertise', Journal of Experimental Psychology: Learning, Memory and Cognition, 10 (1984), 483-495.

Bartos, J. A., The Assessment of Probability Distributions for Future Security Prices, Unpublished Ph.D. thesis, Indiana University, Graduate School of Business, 1969.

Batchelor, R. and Dua, P., 'Forecaster ideology, forecasting technique, and the accuracy of economic forecasts', International Journal of Forecasting, 6 (1990), 3-10.

Bunn, D. and Wright, G., 'Interaction of judgemental and statistical forecasting methods: issues and analysis', Management Science, 37 (1991), 501-518.

Daan, H. and Murphy, A. H., 'Subjective probability forecasting in the Netherlands: some operational and experimental results', Meteorologische Rundschau, 35 (1982), 99-112.
DeBondt, W. F. M., 'What do economists know about the stock market?' Journal of Portfolio Management, 4 (1991), 84-91.

DeBondt, W. F. M. and Thaler, R. M., 'Do security analysts overreact?' American Economic Review, 80 (1990), 52-57.

Donihue, M. R., 'Evaluating the role judgment plays in forecast accuracy', Journal of Forecasting, 12 (1993), 81-92.

Elton, E. J., Gruber, M. J., Das, S. and Hklarka, M., 'Efficiency with costly information: a reinterpretation of evidence from managed portfolios', Unpublished manuscript, New York University, 1991.

Flores, B. E., Olson, D. L. and Wolfe, C., 'Judgmental adjustment of forecasts: a comparison of methods', Interntional Journal of Forecasting, 7 (1992), 421-433.

Gigerenzer, G., Hoffrage, U. and Kleinbolting, H., 'Probabilistic mental models: a Brunswikian theory of confidence', Psychological Review, 98 (1991), 506-528.

Granger, C. W. J., 'Forecasting stock market prices: lessons for forecasters', International Journal of Forecasting, 8 (1992), 3-13.
Granger, C. W. J. and Newbold, P., Forecasting Economic Time Series, 2nd edition, San Diego: Academic Press, 1986.

Griffin, D. and Tversky, A., 'The weighting of evidence and the determinants of confidence', Cognitive Psychology, 24 (1992), 411-435.

Hammond, K. R., 'Probabilistic functionalism: Egon Brunswick's integration of the history, theory and method of psychology', in Hammond, K. R. (ed.), The Psychology of Egon Brunswick, New York: Holt, Rinehart and Winston, 1966.

Ippolito, R. A., 'Efficiency with costly information: a study of mutual fund performance, 1965-1984', Quarterly Journal of Economics, 104 (1989), 1-23.

Jensen, M. C., 'The performance of mutual funds in the period 1945-64', Journal of Finance, 2 (1968), 389-416.

Jensen, M. C., 'Some anomalous evidence regarding market efficiency', Journal of Financial Economics, 6 (1978), 95-101.

Juslin, P., 'An explanation of the hard-easy effect in studies of realism of confidence in one's general knowledge', European Journal of Cognitive Psychology, 5 (1993), 55-71.
Keane, M. P. and Runkle, D. E., 'Testing the rationality of price forecasts: new evidence from panel data', American Economic Review, 80 (1990), 714-735.

Klein, L. R., An Essay on the Theory of Economic Prediction. Chicago: Markham, 1971.

Liljeblom, E., 'An analysis of earnings per share forecasts for stocks listed on the Stockholm Stock Exchange', Scandinavian Journal of Economics (1989), 565-581.

Markowitz, H., Portfolio Selection: Efficient Diversification of Investments, New York: John Wiley, 1959.

McNees, S. K., 'The role of judgment in macroeconomic forecasting accuracy', International Journal of Forecasting, 6 (1990), 287-299.

Murphy, A. H. and Winkler, R. L., 'Subjective probability experiments in meteorology: some preliminary results', Bulletin of the American Meteorological Society, 55 (1974), 1206-1216.

Murphy, A. H. and Winkler, R. L., 'Diagnostic verification of probability forecasts', International Journal of Forecasting, 7 (1992), 435-455.
Murphy, G. L. and Wright, J. C., 'Changes in conceptual structure with

expertise: differences between real world experts and novices', Journal of Experimental Psychology: Learning, Memory and Cognition, 10 (1984), 144-155.

O'Connor, M., 'Models of human behaviour and confidence in judgement: a review', International Journal of Forecasting, 5 (1989), 159-169.

Onkal, D. and Muradoglu, G., 'Evaluating probabilistic forecasts of stock prices in a developing stock market,' European Journal of Opertional Research, 74 (1994), 350-358.

Sengul, G. M. and Onkal, D., 'Semistrong form efficiency in a thin market: a case study'. Paper presented at the 19th Annual Meeting of the European Finance Association, Lisbon, 1992.

Stael von Holstein, C. A. S., 'Probabilistic forecasting: an experiment related to the stock market', Organizational Behavior and Human Performance, 8 (1972), 139-158.

Turner, D. S., 'The role of judgment in macroeconomic forecasting', Journal of Forecasting, 9 (1990), 315-345.

Unal, M., Weak Form Efficiency Tests in Istanbul Stock Exchange, Unpublished MBA Thesis, Bilkent University, Ankara, 1992.

Wolfe, C. and Flores, B., 'Judgmental adjustment of earnings forecasts', Journal of Forecasting, 9 (1990), 83-105.

Wright, G. and Ayton, P., 'The psychology of forecasting', in Wright, G. and Ayton, P. (eds), Judgemental Forecasting, Chichester: John Wiley, 1987, 83-105.

Yates, J. F., 'Analyzing the accuracy of probability judgments for multiple events: an extension of the covariance decomposition', Organizational Behavior and Human Decision Processes, 42 (1988), 281-299.

Yates, J. F., McDaniel, L. S. and Brown, E. S., 'Probabilistic forecasts of stock prices and earnings: the hazards of nascent expertise', Organizational Behavior and Human Decision Processes, 49 (1991), 60-79.

Zarnovitz, V., 'Rational expectations and macroeconomic forecasts', Journal of Business and Economic Statistics, 3 (1985), 293-311.

AUTHORS' BIOGRAPHIES:

Dilek Onkal is an Assistant Professor of Decision Sciences at Bilkent University, Turkey. She received a Ph.D. in Decision Sciences from the University of Minnesota, and is doing research on decision **analysis** and probability forecasting.

Gulnur Muradoglu is an Assistant Professor of **Finance** at Bilkent University, Turkey. She received a Ph.D. in Accounting and Finance from Bogazici University and is doing research on stock market efficiency and stock price forecasting.

AUTHORS' ADDRESSES:

Gulnur Muradoglu and Dilek Onkal, Faculty of Business Administration, Bilkent University, 06533 Ankara, Turkey.

THIS IS THE FULL-TEXT.

Descriptors: Studies; Portfolio management; Forecasting; Stock prices; Statistical analysis; Investment advisors
Classification Codes: 9130 (CN=Experimental/Theoretical); 3400 (CN=Investment analysis); 8130 (CN=Investment services)